**Predictive Model Plan – Student Template**

Use this template to structure your submission. You can copy and paste content from GenAI tools and build around it with your own analysis.

# 1. Model Logic (Generated with GenAI)

Use a GenAI tool (e.g., ChatGPT, Gemini) to generate the logic or structure of your predictive model.  
- You may include pseudo-code, a step-by-step process, or a simplified code snippet.  
- Briefly explain what the model is designed to do.

Paste your GenAI-generated output below or describe the logic in your own words:

The goal is to predict whether a customer will become delinquent based on input features such as income, credit utilization, missed payments, and debt-to-income ratio. Here's a simplified GenAI-generated modeling pipeline: The model ingests cleaned and engineered customer data, transforms it through scaling and encoding, and produces a binary prediction: 0 = Non-delinquent, 1 = Delinquent.

# 2. Justification for Model Choice

Explain why you selected this specific model type (e.g., logistic regression, decision tree, neural network). Consider:  
- Accuracy  
- Transparency  
- Ease of use or implementation  
- Relevance for financial prediction  
- Suitability for Geldium’s business needs

We considered two options:

* **Logistic Regression (Simple):**
  + **Pros:** High interpretability, transparency, fast to train, and easy to explain to business users.
  + **Cons:** May underperform on nonlinear patterns or complex feature interactions.
* **Decision Tree (Complex):**
  + **Pros:** Captures non-linear relationships and feature interactions; intuitive for business users to visualize.
  + **Cons:** Can overfit on small datasets unless pruned; harder to generalize without tuning.

✅ **Recommended Model:** **Decision Tree**

* **Why:** Offers a good balance between interpretability and the ability to capture subtle risk patterns in customer behavior (e.g., multiple missed payments combined with high debt ratio).
* Suitable for financial institutions due to its ability to justify decisions in regulatory settings.

# 3. Evaluation Strategy

Outline how you would evaluate your model’s performance. Include:  
- Which metrics you would use (e.g., accuracy, precision, recall, F1 score, AUC)  
- How you would interpret those metrics  
- Any plans to detect or reduce bias in your model  
- Ethical considerations in making predictions about customer financial behavior

**Key Evaluation Metrics:**

* **Accuracy**: Overall percentage of correct predictions.
* **F1 Score**: Balances precision and recall; essential for imbalanced classes.
* **AUC (Area Under ROC Curve)**: Measures model’s ability to rank positive vs. negative outcomes.
* **Fairness Metrics**: Compare false positive and false negative rates across groups (e.g., by gender, income level).

**Bias Mitigation Techniques:**

* Use **stratified sampling** to balance class representation.
* Apply **SMOTE** (Synthetic Minority Over-sampling Technique) if delinquent cases are underrepresented.
* Conduct **disparate impact analysis** post-training to detect biased outcomes.

**Ethical Considerations:**

* Ensure transparency in feature usage—avoid including race, gender, or proxies.
* Justify automated decisions, especially in loan approvals or credit limits.
* Regularly audit model outputs for drift or unintended bias.